

## Conciencia y dualismo

David Pineda

### ABSTRACT

In this paper I present an objection to the Kripke-Chalmers version of the conceivability argument against physicalism. I also assess the prospects of this physicalist response in the light of two further difficulties: the epistemology of modality and the property dualism argument.

### RESUMEN

En este trabajo presento una objeción a la versión Kripke-Chalmers del argumento concebibilista contra el fisicismo. Asimismo evaluó dicha respuesta fisicista a la luz de otras dos dificultades: la epistemología de la modalidad y el argumento del dualismo de propiedades.

### I. INTRODUCCIÓN

El debate filosófico actual sobre la conciencia está dominado por el argumento concebibilista contra el fisicismo basado en la conciencia fenoménica. Como justificaré en seguida se trata de un argumento de resonancias cartesianas si bien incorpora elementos que provienen de aportaciones de la filosofía reciente, como la semántica bidimensionista. En lo que sigue presentaré una serie de consideraciones sobre este argumento. En esencia, formularé una objeción a la versión Kripke-Chalmers del mismo y evaluaré la posición a la que nos lleva dicha objeción a la luz de otro argumento dualista, el argumento del dualismo de propiedades.

### II. LA TESIS DEL HIATO EXPLICATIVO

Un estado mental es fenoménico-consciente si y sólo si ejemplificarlo conlleva sentirse de un modo determinado. Las experiencias perceptivas, como las experiencias cromáticas, las sensaciones corporales, como el dolor, y

las emociones, como el miedo, son ejemplos paradigmáticos de estados fenoménico-conscientes. No existe un dolor o un miedo neutro, esto es, que no venga acompañado de un particular modo de sentir. De hecho los dolores o los miedos neutros son inconcebibles. El carácter fenoménico de un estado fenoménico-consciente parece pues una propiedad necesaria de dicho estado. Llamaremos a esto la “premisa de Kripke”, pues de hecho juega un papel crucial, como veremos, en el argumento antimaterialista de *Naming and Necessity*.

Por supuesto no todo estado mental es fenoménico-consciente. Por ejemplo, las creencias son perfectamente neutras en el sentido de ‘neutro’ del párrafo anterior. Creer, por ejemplo, que la Tierra gira alrededor del Sol no comporta modo alguno de sentirse característico. Es cierto que a uno tal creencia puede hacerle sentirse mal, como a los científicos píos del siglo XVII a los que creer eso les llamaba a darse cuenta de que la Biblia no era un cúmulo de verdades irrefutables sobre el mundo. Pero ello es debido a que en tales individuos la creencia en cuestión causó cierta emoción negativa la cual sí es fenoménico-consciente, pero no a que la creencia misma no sea neutra. De hecho, hoy en día tal creencia no produce ninguna emoción especial, ni siquiera en los creyentes. Las instituciones eclesiásticas se han encargado, con el correr de los siglos, de romper el vínculo de una creencia tal con emociones negativas.

Los estados fenoménico-conscientes han supuesto un problema filosófico serio al menos desde los tiempos de Descartes. Sucede que tenemos la fuerte intuición de que el carácter fenoménico de tales estados sólo puede ser revelado mediante introspección. Así, suele argumentarse que un ciego de nacimiento no tiene modo de conocer el carácter fenoménico de un estado perceptivo producido por la contemplación de una superficie de color rojo en un ser humano con visión normal. Casi todos nosotros, por ejemplo, nos hemos sentido frustrados al tratar de explicarle el sabor que tiene un ingrediente culinario especial a alguien que no ha probado nada parecido.

La intuición ha tomado cuerpo de tesis filosófica y constituye de hecho una premisa fundamental del argumento concebibilista moderno contra el fisicismo. Podemos llamarla la tesis del carácter esencialmente introspectivo de los estados fenoménico-conscientes (o, siendo más estrictos, del carácter fenoménico de tales estados, aunque en lo sucesivo obviaré esta precisión), esto es, tales estados no pueden ser conocidos por vía no introspectiva. Sin embargo, la tradición filosófica no ha generado otro argumento para dicha importante tesis al margen de describir ciertos experimentos mentales que no hacen sino elaborar la intuición anterior. El más antiguo de estos experimentos es el del espectro invertido, de raíces lockeanas. Parece concebible alguien con un comportamiento e incluso organización funcional idéntica a la de un ser humano normal que sin embargo experimente los colores de un modo invertido al nuestro, esto es, que en él las superficies rojas generen el

tipo de experiencias cromáticas que al resto nos generan las superficies verdes y a la inversa.

En efecto, que un sujeto tal parezca concebible tiene mucho que ver con la tesis del carácter esencialmente introspectivo de los estados fenoménico-conscientes. Esto es, aunque dispongamos de todo tipo de información sobre ese sujeto adquirida por vía no introspectiva, esa información no nos revelará el carácter fenoménico de sus experiencias perceptivas, precisamente porque tal carácter es necesariamente introspectivo. Esta es la conclusión a la que pretende llevarnos otro famoso experimento mental, éste mucho más reciente, el argumento del conocimiento [Jackson (1982)]. Se trata de imaginar a una mujer, Mary, confinada por mor del ejemplo en una habitación donde todo es blanco y negro y cuya conexión con el exterior se lleva a cabo mediante un ordenador potente cuya pantalla también es en blanco y negro. Sucede que Mary es una chica muy lista, se interesa por la percepción humana, y con el tiempo llega a saberlo todo acerca de la neurofisiología de la percepción humana del color. Todo ese saber, no obstante, supuestamente no le permitiría a Mary acceder al carácter fenoménico de las experiencias cromáticas humanas. Sólo cuando es liberada de la prisión y experimenta por primera vez algo de color rojo, digamos un puñado de fresas, accede a ese conocimiento.

No hay otro modo de acceder a ese conocimiento, se quiere concluir, esto es, la tesis del carácter esencialmente introspectivo de las experiencias cromáticas. Mary conoce perfectamente qué sucede en un cerebro humano cuando está percibiendo una superficie de color rojo, pero ello no la lleva a conocer el carácter fenoménico de tal tipo de percepción. La tesis del hiato epistémico de lo fenoménico (término acuñado en Levine (1983) que ha hecho fortuna) consiste en generalizar este resultado para cualquier tipo de estado fenoménico-consciente. Esto es, sostiene que no hay modo de explicar físicamente, es decir, usando conceptos físicos, el carácter fenoménico de tales estados. Existe un salto epistémico, un hiato invadable, entre lo fenoménico y lo físico.

Nótese que la tesis del hiato explicativo es más débil que la del carácter esencialmente introspectivo de los estados fenoménico-conscientes y, de hecho, la segunda implica la primera bajo ciertos supuestos. La tesis del hiato se limita a sostener que existe este salto epistémico entre lo fenoménico y lo físico mientras que la otra proporciona una explicación de por qué ocurre ese salto en realidad. La razón sería justamente que el carácter fenoménico de un estado mental sólo es accesible por vía introspectiva, mientras que cualquier explicación física usará conceptos físicos que no se adquieren por esa vía. De ahí que una explicación tal dejará siempre fuera ese elemento fenoménico.

Aquí es necesario precisar algo mejor la tesis del hiato. La hemos formulado como la tesis de que existe un salto epistémico entre lo fenoménico y lo físico, ¿pero se sigue de ello que lo fenoménico no es físico? Si parece que sí a partir de nuestra formulación entonces deberíamos mejorarla. O al menos

eso piensan muchos filósofos que llamaré “compatibilistas” (para algunos trabajos recientes compatibilistas véanse Loar (1997) y (2003), Levine (1998), Levin (2002), Block (2007)). Los filósofos compatibilistas son fisicistas y, en consecuencia, sostienen que lo fenoménico es físico, como todo lo demás que existe en el mundo real (empírico), pero sostienen también que la tesis del hiato es correcta, esto es, que dicha tesis es *compatible* con el fisicismo. Su posición se resume en que si bien hay un salto epistémico entre lo físico y lo fenoménico, esto es, no es posible ofrecer una explicación de lo fenoménico usando sólo conceptos físicos, no hay sin embargo un salto metafísico. El salto epistémico se debería en todo caso, según este punto de vista, a la distinta naturaleza de los *conceptos* fenoménicos con respecto a los conceptos físicos, pero no al hecho de que tales conceptos denoten entidades de naturaleza metafísica distinta.

Más adelanté volveré a los filósofos compatibilistas, pero traerles a colación aquí sirve para llamar la atención de que la tesis del hiato explicativo por sí sola no parece bastar para argumentar una posición contraria al fisicismo. El filósofo dualista debe argumentar, en efecto, que el salto o hiato epistémico entre lo fenoménico y lo físico obedece a un salto ontológico, pues, de hecho, el fisicismo es una tesis metafísica, no epistemológica. Se encuentra entonces en una situación dialéctica muy parecida a la de Descartes en las *Meditaciones*. En efecto, mediante hipótesis escépticas radicales como la del genio maligno, el filósofo francés pretende haber establecido que puedo concebirme con claridad y distinción como ente pensante desprovisto de cuerpo. Añade entonces la premisa crucial de que las cosas pueden ser tal y como las concibo clara y distintamente. De ambas premisas se sigue que yo en tanto que ente pensante no soy mi cuerpo, pues puedo existir sin él.

La segunda es una premisa racionalista. Sostiene que lo que yo puedo concebir *a priori* establece qué es metafísicamente posible, de ahí que si yo puedo concebir (de modo claro y distinto) mi mente sin el cuerpo se sigue que mi mente puede existir sin el cuerpo y por tanto es una sustancia distinta de éste.

De modo similar, el dualista moderno parte de la tesis del hiato explicativo. Dado cualquier correlato físico de un estado fenoménico establecido empíricamente, por ejemplo, la oscilación córtico-talámica en el caso del dolor, puedo perfectamente concebir el dolor en ausencia de dicha oscilación y viceversa. Pero si la premisa racionalista es cierta, entonces es posible que se dé el dolor sin oscilación y la oscilación sin dolor, con lo cual el dolor no es la oscilación. La tesis del hiato garantiza que nos encontraremos en una situación similar con respecto a cualquier otro correlato físico del dolor que emane de cualquier propuesta explicación física del mismo, pues una explicación tal será necesariamente incompleta.

## III. EL GIRO KRIPKEANO

En *Naming and Necessity*, Saul Kripke argumentó muy persuasivamente que existen verdades necesarias *a posteriori* y verdades contingentes *a priori*. Ejemplo de las primeras serían identidades teóricas del tipo ‘agua = H<sub>2</sub>O’, ‘sal marina = NaCl’ o bien enunciados que predicen propiedades esenciales de objetos empíricos, por ejemplo, ‘Kripke proviene de la unión de la célula femenina X con la célula masculina Y’.

A primera vista, la existencia de verdades necesarias *a posteriori* parece contravenir la premisa racionalista. Si ‘agua = H<sub>2</sub>O’ es *a posteriori* ello parece querer decir que podemos concebir que el agua no sea H<sub>2</sub>O. Pero si por otra parte es (metafísicamente) necesario, entonces estamos concibiendo algo metafísicamente imposible. Ahora bien, si eso es así, si a veces concebimos imposibilidades metafísicas, tal puede ser también el caso con las situaciones concebibles que emanan de la tesis del hiato. Kripke parece pues estar apuntalando el compatibilismo y serrando una pata crucial al argumento dualista concebilista de corte cartesiano.

Sin embargo, algo sorprendentemente Kripke usará de hecho las necesidades *a posteriori* para reforzar el argumento concebilista. Kripke rechazará que la explicación del carácter *a posteriori* de una identidad como ‘agua = H<sub>2</sub>O’ radique en que podemos concebir una imposibilidad metafísica, a saber, que el agua no sea H<sub>2</sub>O, y ofrecerá una explicación alternativa que de hecho mantiene la premisa racionalista. La explicación de Kripke tiene que ver con importantes tesis semánticas de *Naming and Necessity*. En esencia, la idea es que un término como ‘agua’ expresa un concepto que denota una sustancia “rígidamente”, esto es, tiene la misma denotación en todo mundo posible. Pero el modo como el concepto fija su referencia es mediante ciertos indicios de su referente, en este caso, el agua. Tales indicios, su aspecto para un ser humano o el hecho de que llene los ríos y lagos del planeta, son propiedades del agua. Así, el concepto denota rígidamente aquella sustancia que de hecho responde a esos indicios y que no es otra cosa que el agua.

Ahora bien, en el caso de ‘agua’ resulta que los indicios en cuestión son propiedades *contingentes* del agua. Ciertamente no es una propiedad necesaria del agua llenar los lagos del planeta, como tampoco lo es el impacto que produce en seres con nuestro aparato sensorial. Por razón de ello, sí existen mundos metafísicamente posibles en los que algo que no es agua resulta ser la sustancia que responde a esos indicios. Y eso es exactamente lo que concebimos cuando pensamos en la posible falsedad de ‘agua = H<sub>2</sub>O’, esto es, concebimos una situación genuinamente posible en la que una sustancia diferente del agua resulta tener los indicios que fijan la denotación de nuestro concepto de agua. Pero dado que, de hecho, ha resultado que la sustancia que en efecto responde a esos indicios es el agua y no otra cosa, nuestro concepto de agua denota rígidamente el agua y no otra cosa. La situación genuinamen-

te posible que concebimos es una en la que resulta falsa la identidad ‘la sustancia que llena los lagos del planeta y responde a tales y cuales indicios =  $H_2O$ ’ y no una en la que resulta falsa la identidad ‘agua =  $H_2O$ ’. La primera identidad es contingente, la segunda es necesaria. Ahora bien, dado que los indicios en cuestión fijan la denotación de ‘agua’ la contingencia de la primera identidad explica el carácter *a posteriori* de la segunda. En efecto, necesito información empírica para determinar si el mundo real está entre las situaciones posibles en las cuales la primera identidad es verdadera o bien entre aquellas en las que es falsa, esto es, necesito información empírica para determinar si ‘agua =  $H_2O$ ’ es (necesariamente) verdadera o (necesariamente) falsa.

El giro kripkeano consiste en argumentar que para el caso de una identidad como ‘dolor = oscilación córtico-talámica’ no podemos ofrecer una explicación análoga de su carácter *a posteriori*. Aquí la premisa crucial es lo que he llamado “la premisa de Kripke” al principio de la sección precedente, esto es, nuestro concepto de dolor fija su denotación (un determinado tipo de estado fenoménico-consciente) mediante indicios, el modo característico en que siento el dolor, que son propiedades *necesarias*, no contingentes, del dolor. Según esto, una situación en la que algo distinto de la oscilación córtico-talámica tiene los indicios fenoménicos del dolor es en efecto una situación posible en la cual la oscilación córtico-talámica no es el dolor. Y dado que si ‘dolor = oscilación córtico-talámica’ es verdadero entonces es necesariamente verdadero, se sigue de que haya situaciones posibles de este tipo que de hecho la identidad en cuestión es falsa. Como el argumento puede reproducirse para cualquier otro concepto físico que expresemos a la derecha de la igualdad, obtenemos la conclusión dualista.

Dado que me referiré a ella en lo sucesivo resumiré a continuación la explicación kripkeana de la necesidad *a posteriori*. Una identidad necesaria es *a posteriori* sólo cuando uno de los términos implicados en ella es un designador rígido que fija su denotación mediante indicios que son sólo propiedades contingentes de la entidad denotada. La aposterioridad de una identidad necesaria se explica entonces porque hay situaciones genuinamente posibles en las que los indicios en cuestión responden a entidades que difieren de la denotada rigidamente por el designador.

#### IV. UNA OBJECIÓN A LA EXPLICACIÓN KRIPKEANA

##### IV.1 *La semántica de Boyd*

Al escribir *Naming and Necessity* Kripke tuvo en mente, entre otros ejemplos, lo que hemos llamado identidades teóricas. Esto es, enunciados que consisten en un signo de igualdad flanqueado por dos términos uno de los

cuales es de uso común mientras que el otro es un término teórico, introducido por alguna teoría científica. Nuestro ejemplo ha sido 'agua = H<sub>2</sub>O'.

Richard Boyd, en Boyd (1980), formuló una de las primeras críticas al argumento dualista de Kripke al señalar que el argumento había dejado al menos un cabo suelto. En efecto, Kripke explica la aposterioridad de los enunciados de identidad a partir de importantes tesis semánticas sobre términos de uso común como 'agua' o 'dolor', pero no dice una palabra sobre términos teóricos como 'H<sub>2</sub>O' o 'oscilación córtico-talámica'. Pero esta omisión no es inocente, porque de hecho, para que el argumento concluya, es necesario suponer que los términos teóricos funcionan como 'dolor' y no como 'agua', esto es, son designadores rígidos que fijan su denotación mediante indicios que son necesarios y no meramente contingentes de la entidad designada. Pues si 'oscilación córtico-talámica' funciona como 'agua' y no como 'dolor' entonces podemos, contrariamente a lo que supone Kripke, ofrecer una explicación kripkeana de la aposterioridad de 'dolor = oscilación córtico-talámica' compatible con su verdad: a saber, una que invoca una situación genuinamente posible en la cual algo que no es la oscilación córtico-talámica, y por tanto no es el dolor, posee los indicios que fijan la denotación de 'oscilación córtico-talámica'.

Esto es exactamente lo que propuso Boyd como respuesta a Kripke en el trabajo aludido; supuso, usando una teoría de inspiración funcionalista, que los términos teóricos funcionan análogamente a 'agua' y no a 'dolor'. Dicho de otro modo, sostuvo que los términos teóricos son designadores rígidos que fijan su denotación mediante indicios que son sólo propiedades contingentes de la entidad que designan. Es interesante señalar que esta semántica de Boyd sobre términos teóricos es perfectamente compatible con la semántica de Kripke para términos de género natural, con la explicación kripkeana de la necesidad *a posteriori* e incluso con la premisa racionalista del argumento concebibilista. Lo único que hace es llenar un hueco en la semántica de Kripke, pero lo llena de tal modo que en realidad desactiva la versión kripkeana del argumento concebibilista. De hecho la posición de Boyd desactiva también el argumento concebibilista general descrito en la sección segunda. Recordemos que ese argumento tomaba como premisas la tesis del hiato explicativo y la premisa racionalista que la concebibilidad es nuestra guía a la posibilidad metafísica. Sin embargo, dada la semántica de Boyd puede defenderse que no podemos concebir que el dolor no sea oscilación córtico-talámica, pero sí podemos concebir, sí es genuinamente posible, que algo que no es el dolor tenga los indicios contingentes que fijan la denotación de 'oscilación córtico-talámica'. Esa genuina posibilidad contribuiría a explicar el por qué de la tesis del hiato, de un modo compatible con el fisicismo. El resto de la explicación apelaría al carácter único de conceptos fenoménicos como el expresado por el término 'dolor', que según la semántica de Boyd, y a diferencia de la de Kripke, devendrían únicos en su especie.

En la más reciente y detallada defensa de su propia versión del argumento concebibilista [Chalmers (2007)], David Chalmers dice suscribir la tesis de Boyd sobre los términos teóricos y reconoce de modo consecuente que tal tesis plantea una salida fisicista al argumento concebibilista. Sin embargo, la desestima con el argumento de que la posición de Boyd tiene como consecuencia que la conciencia fenoménica es un rasgo físico del mundo inaccesible desde los conceptos físicos y por ello opaco a cualquier explicación física. En efecto, recordemos que según la semántica de Boyd los términos teóricos en general, aquellos en los que se articularía una posible explicación científica del mundo, identifican una entidad física sólo a través de indicios contingentes de la misma. Por ello tales conceptos, exactamente como ocurre con el concepto expresado por el término ‘agua’, pueden a lo sumo ser detectores fiables de la entidad designada, en un entorno determinado, pero desde luego no revelan su naturaleza.

Chalmers denomina “monismo russelliano” a la posición fisicista a la que aboca la semántica de Boyd de los términos teóricos, debido a que en algún momento Russell pareció defender una versión del materialismo acorde con esta tesis. Sin embargo, en mi opinión la semántica de Boyd tiene consecuencias poco digeribles que van más allá del problema suscitado por la conciencia fenoménica e incluso de la cuestión del fisicismo. Si los conceptos científicos jamás revelan la naturaleza de las entidades que denotan sino tan sólo indicios contingentes, entonces tal naturaleza deviene inaccesible a la ciencia humana, lo que equivale a decir inaccesible a los esfuerzos humanos de conocimiento más sistemáticos y desarrollados. En otros términos, la propuesta de Boyd parece llevarnos a una ontología kantiana, según la cual la naturaleza de las cosas es una especie de “noumenon”, opaco al conocimiento humano.

#### IV.2 *Enunciados de reducción interteórica*

Resumamos los resultados alcanzados en la discusión de la sección anterior. Por un lado, tenemos la versión Kripke-Chalmers del argumento concebibilista que nos lleva a una posición dualista; por otro, tenemos la semántica de Boyd sobre los términos teóricos que desactiva el argumento pero nos lleva ineludiblemente a una tesis ontológica kantiana sobre el mundo. Ambas alternativas parecen poco deseables para el fisicista. En esta sección, no obstante, ofreceré algunas razones para pensar que existe alguna alternativa mejor.

Dije al principio de la sección anterior que en *Naming and Necessity* Kripke tuvo claramente en mente ejemplos de identidades teóricas. En este tipo de identidades, uno de los términos que flanquea el signo de identidad es de uso corriente, mientras que el otro es un término teórico. Sin embargo,

existen otro tipo de identidades cuyos dos términos son teóricos, es el caso de enunciados de reducción interteórica. Un ejemplo (esquemático) podría ser:

$$\text{'H}_2\text{O} = \text{Q'}$$

donde Q es una descripción mecánico-cuántica del agua. Dando por buena la semántica de Kripke de términos de género natural, según la cual son designadores rígidos, tal identidad resulta ser necesaria. Ahora bien, también parece razonable suponer que se trata de un enunciado *a posteriori*, pues expresa la identidad de *designata* entre términos introducidos por teorías científicas contrastadas empíricamente de modo independiente. En suma, los enunciados de reducción interteórica son ejemplos de necesidades *a posteriori*.

Nótese por de pronto que la semántica de Boyd nos proporciona una explicación kripkeana de la aposterioridad de este tipo de enunciados necesarios, lo cual refuerza dicha propuesta, pues ahora ya no aparece meramente como una opción *ad hoc* para sortear la dificultad que presenta el argumento concebibilista. Aunque él no lo dice explícitamente, sospecho que esta es de hecho una de las razones por las que Chalmers ha decidido suscribir la semántica de Boyd.

Pero antes hemos visto los inconvenientes de la semántica de Boyd. ¿Existe alguna alternativa? Bien, la alternativa más obvia a que un término teórico fije su denotación por medio de indicios contingentes de la misma es que la fije por medio de indicios necesarios, es decir, de propiedades necesarias de la entidad que denota. ¿Pero no supone esto volver a la semántica de los términos teóricos tácitamente supuesta en *Naming and Necessity*, a saber, que son términos que funcionan como 'dolor'? ¿Y no supone esto a su vez la imposibilidad de obtener una explicación kripkeana de los enunciados de reducción interteórica?

El físico preocupado por el argumento concebibilista, o por la versión del mismo de Kripke-Chalmers, desde luego tiene la fuerte tentación de responder afirmativamente en ambos casos. En efecto, responder afirmativamente a ambas cuestiones equivale a rechazar lo que al final de la sección tercera he llamado la explicación kripkeana de la necesidad *a posteriori*, es decir, defender que no toda necesidad *a posteriori* puede explicarse kripkeamente. Ahora bien, dado que esa explicación es una premisa crucial en la versión Kripke-Chalmers del argumento concebibilista, dicho argumento no concluye.

Pero haciendo justicia al dualista es necesario reconocer que existen otras alternativas con respecto a la semántica de los términos teóricos al margen de la de Boyd y la del propio Kripke. Hasta ahora nos hemos referido de modo general a propiedades necesarias de una entidad, sin hacer ulteriores distinciones. Pero hay, de hecho, importantes distinciones que hacer, y es importante examinarlas para nuestra discusión.

En primer lugar, Kit Fine ha dado ejemplos de propiedades necesarias no esenciales. Así, ser el único elemento del conjunto unitario de Sócrates es una propiedad necesaria de Sócrates, una propiedad que Sócrates tiene en todo mundo posible en el que existe, pero ciertamente no diríamos que es parte de la esencia o naturaleza de Sócrates [Fine (1994)]. Ello motiva una distinción entre propiedades necesarias y propiedades esenciales, si bien dicha distinción no ha podido ser recogida en algo así como una definición del segundo tipo de propiedades.

En segundo lugar, dentro de las propiedades esenciales hay también una importante distinción que hacer. Así, ser mamífero parece ser una propiedad esencial del gato de mi vecina, pero es un tipo de esencia diferente de la que es para el dolor su contenido fenomenológico, según *Naming and Necessity*. En efecto, Kripke dice, y es una premisa crucial en su argumento dualista como recordé al principio de la sección segunda, que nada puede sentirse como el dolor y no ser el dolor. No podemos decir en cambio que nada puede ser mamífero y no ser el gato de mi vecina. El perro de mi otro vecino atestigua esto perfectamente. E incluso aunque el gato de mi vecina fuera el único mamífero del mundo, habría no obstante situaciones posibles en que cosas que no son el gato de mi vecina ejemplificarían la propiedad de ser mamífero. Para disponer de términos que faciliten la discusión, llamaré “esencias individuativas” a las propiedades esenciales del estilo de las propiedades fenomenológicas del dolor y “esencias parciales” a las del estilo de ser un mamífero para el gato de mi vecina. Las siguientes cláusulas formales pretenden aclarar la diferencia entre ellas:

(EP) P es una *esencia parcial* de X sólo si  $\Box$  (si X existe, entonces X tiene P).

(EI) P es una *esencia individuativa* de X sólo si P es una esencia parcial de X y  $\Box \forall y$  (si y tiene P, entonces  $y = X$ ).

Desde luego, la distinción de Fine entre propiedades necesarias y propiedades esenciales no va a ayudarnos con respecto a la amenaza que supone comprometerse con un mundo ontológicamente kantiano. Sin embargo, la distinción entre esencias parciales y esencias individuativas ofrece mejores perspectivas al dualista. Véamoslo.

Por de pronto ahora podemos afirmar que la semántica para términos teóricos tácitamente supuesta en *Naming and Necessity* corresponde a la tesis de que dichos términos fijan su denotación mediante indicios que son esencias individuativas de la entidad que denotan. Pero en realidad, esta vía le está vedada al dualista pues conlleva como hemos visto que no hay explicación kripkeana del estatus de los enunciados de reducción interteórica.

Ahora bien, nuestra distinción entre esencias individuativas y esencias parciales sugiere otras alternativas más prometedoras para el dualista. Plan-

teemos, por ejemplo, la tesis tentativa de que los términos teóricos fijan su denotación mediante indicios que son esencias parciales *pero no individuativas* de las entidades que denotan. Veamos en efecto, en primer lugar, que bajo esta tesis es posible ofrecer una explicación parecida a la kripkeana del estatus de los enunciados de identidad interteórica.

Supongamos que 'H<sub>2</sub>O' denota el agua bajo cierta esencia parcial del agua EPH<sub>2</sub>O mientras que 'Q' denota el agua bajo otra esencia parcial del agua EPQ. Entonces podemos explicar el carácter *a posteriori* de 'H<sub>2</sub>O = Q' del siguiente modo: existe una situación posible en la que hay algo X que tiene EPH<sub>2</sub>O y hay algo Y que tiene EPQ y sin embargo X ≠ Y. La idea es que en esa situación posible ni X ni Y son el agua. Que existe una situación posible así parece garantizado a partir del hecho de que tanto EPH<sub>2</sub>O como EPQ son esencias parciales pero no individuativas del agua.

Es cierto que la explicación propuesta no concuerda con la letra de la explicación kripkeana de la necesidad *a posteriori* tal y como la he formulado al final de la sección segunda, pero ciertamente sí concuerda con su espíritu, pues a fin de cuentas la explicación se basa en que el enunciado 'la entidad que tiene EPH<sub>2</sub>O = la entidad que tiene EPQ' es genuinamente contingente.

Ahora bien, si los indicios mediante los cuales los términos teóricos fijan su denotación son meras esencias parciales, y por tanto hay situaciones posibles en los que otras entidades son las que responden a esos indicios, ¿no existe entonces un modo kripkeano de explicar la identidad psicofísica 'dolor = oscilación córtico-talámica', con lo cual el argumento concebilista queda nuevamente desactivado? La respuesta es negativa, y lo es porque 'dolor' fija su denotación mediante una esencia individuativa del dolor, su contenido fenomenológico. Llamemos a esos indicios EIDOL y a los indicios asociados a 'oscilación córtico-talámica', una supuesta esencia parcial del dolor, EPDOL. No hay entonces una situación posible en que existan cosas distintas X e Y que no sean el dolor y tengan respectivamente EIDOL y EPDOL. No pueda haberla si EIDOL es una esencia individuativa del dolor.

Del razonamiento del párrafo anterior se sigue que una identidad tal que uno de sus términos fije su denotación mediante una esencia parcial y el otro mediante una esencia individuativa de la entidad designada no puede explicarse kripkeanamente, tal como hemos visto que sucede cuando ambos términos fijan su denotación mediante esencias individuativas. Así pues, la opción intermedia según la cual sólo algunos términos teóricos involucran esencias individuativas de la entidad designada mientras que otros involucran tan sólo esencias parciales, está también vedada al dualista partidario del argumento concebilista, pues acarrea la consecuencia de que ciertos enunciados de reducción interteórica no pueden ser explicados kripkeanamente.

La opción buena para el dualista, por lo que llevamos visto hasta ahora, es la de que los términos teóricos fijan su denotación mediante esencias parciales de los *designata*. Sin embargo, esta opción no acaba de evitar la ame-

naza de una ontología kantiana. Ciertamente, si todo lo que la ciencia pudiera revelarnos acerca de un gato es que pertenece a la clase de los mamíferos desde luego gran parte de su naturaleza permanecería oculta, en particular aquella parte que distingue al gato de otros mamíferos. Similarmente, si la ciencia sólo puede darnos esencias parciales del agua, entonces la naturaleza del agua no nos es totalmente revelada a través de ella, sigue habiendo una parte oculta, un “noumenon” inaccesible. En definitiva, la opción “buena” parece adolecer del mismo problema que la semántica de Boyd, con lo que nuestra conclusión provisional debe ser que el estatus de los enunciados de reducción interteórica es después de todo un buen antídoto fisicista contra la versión Kripke-Chalmers del argumento dualista concebibilista.

#### V. RACIONALISMO MODAL Y DUALISMO DE PROPIEDADES

Es el momento de evaluar las consecuencias de la posición fisicista a la que nos ha llevado la discusión del argumento Kripke-Chalmers, en nuestro intento de presentar una defensa razonable ante tal argumento.

Por de pronto, de nuevo hay que hacer justicia al dualista y reconocer que existen todavía dos alternativas que no hemos examinado. Una de ellas dice que todos los términos científicos para el agua revelan tan sólo una esencia parcial de esta sustancia pero que la unión de todas ellas configura una (la) esencia individuativa. De este modo, se evita la ontología kantiana y al mismo tiempo se retienen las virtudes (para el dualista) de la hipótesis examinada al final de la sección anterior. Un problema, entre otros, de esta alternativa es que parece incompatible con la eventual infinita división de la materia, eventualidad que algunos físicos contemporáneos no desdeñan en absoluto. Si la eventualidad llegara a cumplirse, entonces habría una sucesión sin fin de términos científicos para el agua, cada uno revelando tan sólo una esencia parcial, introducidos por teorías científicas que tratan sobre cada uno de los infinitos niveles de la materia. Y desde luego, la esencia individuativa del agua, supuestamente obtenida de la unión de esta serie infinita de esencias parciales sería de nuevo humanamente inalcanzable.

La otra posibilidad es sugerir que sólo uno de los términos científicos del agua revela una esencia individuativa mientras que el resto opera por medio de indicios contingentes. Veamos cómo, en efecto, podemos explicar kripkeanamente la identidad ‘ $H_2O = Q$ ’ bajo este supuesto. Supongamos que la esencia individuativa del agua que fija la denotación de ‘ $Q$ ’ es EI $Q$  mientras que la propiedad contingente del agua que fija la denotación de ‘ $H_2O$ ’ es CONT. Pues bien, en ese caso hay una situación posible en la que algo  $X$ , a saber el agua, tiene EI $Q$  y algo  $Y$  que no es el agua tiene CONT. Basta pues con suponer una situación posible en la que hay agua pero no tiene la propiedad CONT sino que de hecho la tiene alguna otra sustancia  $Y$ . Que hay una

situación posible tal parece de nuevo garantizado por el hecho de que CONT es sólo una propiedad contingente del agua. Por otro lado, una identidad que apareje el concepto de dolor con un concepto teórico que opere mediante una esencia individuativa tampoco podría explicarse kripkeamente.

Ahora bien, es importante no perder de vista la dialéctica en la que estamos inmersos. El dualista está tratando de convencernos de la falsedad del fisicismo por medio de una versión del argumento concebibilista. Pero nuestro razonamiento nos lleva a la conclusión de que tal argumento sólo funciona si nos comprometemos con determinada semántica de los términos teóricos. ¿Y por qué debería el fisicista comprometerse con una semántica tal? ¿O comprometerse con el rechazo a la tesis de la infinita divisibilidad de la materia? El argumento dualista, así pues, parece irremediablemente cojo a menos que esa semántica pueda justificarse independientemente.

### V.1 *El racionalismo modal*

Una posible justificación independiente es ésta. La semántica de los términos teóricos que debe proponer el dualista tiene la virtud de permitir ofrecer una explicación kripkeana (al menos en espíritu) de la necesidad *a posteriori*. Y a su vez dicha explicación tiene la virtud de respetar el racionalismo modal, esto es, la tesis de que tenemos un acceso *a priori* de las posibilidades, que nuestro conocimiento modal tiene un fundamento *a priori*.

El racionalismo modal es valioso por cuanto ofrece una explicación razonable del conocimiento de verdades modales. Su antagonista, la teoría a la que debería tender un fisicista que quiera ofrecer ante el argumento concebibilista el tipo de respuesta que hemos esbozado en el apartado anterior, es el empirismo modal. Esto es, la tesis de que las verdades modales se justifican en todo caso *a posteriori* y son por tanto conocidas en base a datos y generalizaciones de carácter empírico. La dificultad obvia a la que se enfrenta este tipo de epistemología modal, y que elude por su parte el racionalismo, es la histórica observación humeana de que la experiencia sólo puede darnos a conocer cómo son de hecho las cosas, pero no cómo deben de ser. En consecuencia, el fisicista en cuestión parece tener sólo dos opciones abiertas: o bien desarrollar una epistemología modal racionalista alternativa a la que lleva a la explicación kripkeana y que permita dar cuenta de todos los casos paradigmáticos de necesidad *a posteriori*; o bien encontrar un modo razonable de bloquear la crítica humeana.

### 5.2 *El argumento del dualismo de propiedades*

Otro punto a tener en cuenta proviene de traer a colación otro argumento dualista basado en la conciencia fenoménica, un argumento menos profusamente

comentado y discutido que el argumento concebibilista, pero a mi juicio tanto o incluso más problemático para el fisicismo.

Dicho argumento, formulado y discutido en White (1983), proviene en realidad de una objeción de Max Black a la teoría de la identidad psiconeural defendida en los años cincuenta por Smart y Place [véase Block (2007)]. Tiene por tanto también antecedentes añejos, si bien no tanto como el argumento concebibilista. Siguiendo la terminología de White, las premisas básicas del argumento son ciertas tesis sobre los conceptos que denomina “fregeanismo mínimo”:

- (1) Dos conceptos codenotativos distintos involucran modos de presentación distintos del mismo referente.
- (2) Un modo de presentación de un referente involucra esencialmente una propiedad ejemplificada por el referente.

A la vista de (1) y (2) se entiende la denominación de White. Los modos de presentación en cuestión son cierta versión de los sentidos fregeanos. Consideremos ahora un enunciado de identidad psicofísica, como nuestro recurrente ‘dolor = oscilación córtico-talámica’. Dado que este enunciado es *a posteriori* y no analítico es claro que ‘dolor’ y ‘oscilación córtico-talámica’ expresan conceptos distintos. Bajo el supuesto de que la identidad es verdadera, no obstante, cabe concluir que esos dos conceptos son codenotativos. En aplicación de las premisas (1) y (2), por tanto, cada uno involucra una propiedad diferente de esa misma entidad denotada por ambos.

Lo interesante del argumento es el siguiente paso. Si el fisicista tiene razón, las propiedades involucradas en ambos conceptos deben ser físicas o bien propiedades neutras como las propiedades funcionales. Digamos para abreviar que deben ser propiedades físico-funcionales. Ahora bien, supongamos pues que la propiedad que individualiza el modo de presentación fenoménico expresado por el término ‘dolor’ es una propiedad físico-funcional F. Una propiedad así debe poder ser descrita por conceptos físico-funcionales. Llamemos ‘F’ al término que expresa tal descripción. Ahora bien, bajo todos esos supuestos, el enunciado (\*) ‘el dolor es el estado que tiene F’ debe valer *a priori*, por razones parecidas a aquellas por las que según el kripkeano el enunciado ‘el agua es la sustancia que llena los ríos y lagos de nuestro entorno’ debe también valer *a priori*.

No obstante, y esta es la dificultad para el fisicista, se desprende de la tesis del hiato explicativo que enunciados como (\*) no pueden valer *a priori*, pues de lo contrario la implicación:

X tiene F  


---

X es dolor

valdrá también *a priori* y, en consecuencia, una explicación del dolor en términos físico-funcionales devendrá completa, sin huecos, algo que contradice frontalmente la tesis del hiato. En definitiva, lo que parece mostrar el argumento del dualismo de propiedades es que las tesis fregeanas (1) y (2) sobre los conceptos, la tesis del hiato explicativo y el fisicismo forman un conjunto inconsistente de premisas. En consecuencia, un filósofo compatibilista, como aquellos a los que aludí al final de la sección segunda, es decir, un filósofo comprometido con la segunda y la tercera tesis, debe argumentar contra el fregeanismo mínimo.

Y la tarea es ardua, porque exige nada menos que desarrollar una teoría de los conceptos alternativa al fregeanismo mínimo que dé cuenta, además, del supuesto fenómeno del hiato explicativo. Por de pronto veamos que el fregeanismo mínimo no puede desbaratarse simplemente tomando en consideración enunciados de reducción interteórica, como propusimos en el caso del argumento concebibilista de Kripke-Chalmers. Así, el carácter *a posteriori* de ' $H_2O = Q$ ' podría explicarse en principio suponiendo simplemente que los conceptos expresados por ' $H_2O$ ' y ' $Q$ ' involucran modos de presentación individuados por propiedades físicas empíricas distintas del agua. Mientras que su carácter necesario podría explicarse siguiendo la doctrina kripkeana de los términos teóricos de géneros naturales como designadores rígidos. En efecto, las premisas (1) y (2) no son por sí mismas incompatibles con la semántica kripkeana, de ahí que se hable tan sólo de fregeanismo "mínimo".

De hecho, los filósofos compatibilistas han tratado de objetar a aquello mínimamente fregeano que se encuentra en las tesis (1) y (2). Posiblemente la teoría compatibilista más influyente y desarrollada sea la de Brian Loar [Loar (1997), (2003)]. Loar sostiene que los conceptos fenoménicos como aquél expresado por el término 'dolor' son conceptos reconocitivos demostrativos-tipo. No es necesario entrar en los detalles de la propuesta de Loar, baste con decir que según dicha teoría no hay nada parecido a un modo de presentación en el caso de los conceptos fenoménicos, sino que su vinculación con la entidad que designan es directa.

Naturalmente, si no hay una propiedad del estado de dolor que medie entre el concepto fenoménico y dicho estado para propiciar la referencia, entonces el argumento del dualismo de propiedades no puede despegar. No es necesario comprometerse con propiedades de naturaleza no físico-funcional. Ahora bien, si los conceptos fenoménicos carecen de cualquier grosor cognitivo, como parece desprenderse de esta propuesta, entonces pende una explicación de por qué existe un hiato explicativo entre lo fenoménico y lo físico.

Joseph Levine ha expresado la dificultad que afronta aquí el compatibilista de modo muy gráfico con el siguiente ejemplo [Levine (2001)]. Supongamos que entro en una habitación con los ojos vendados y señalo hacia algún punto. Si a mi pregunta "¿qué estoy señalando?" alguien que asiste a la escena respondiera "una lámpara de pie", pues eso es, en efecto, aquello a lo

que estoy señalando a ciegas, entonces resultaría incomprensible que mi reacción ante tal respuesta fuera: “¿y por qué es una lámpara de pie aquello a lo que estoy señalando, y no otra cosa?; tu explicación no me satisface”. De modo similar, si los conceptos fenoménicos como el expresado por ‘dolor’ funcionan como “demostrativos ciegos”, por seguir la sugerente terminología de Levine, entonces dada la justificación empírica de que aquello a que apunta tal concepto es un estado físico, digamos, oscilación córtico-talámica, resultaría igualmente incomprensible preguntarse por qué dicho concepto apunta a tal estado físico en lugar de a cualquier otra cosa.

La observación de Levine parece dejar claro que los conceptos fenoménicos no pueden funcionar como demostrativos ciegos según el filósofo compatibilista que quiere a la vez reconocer un hiato explicativo para la conciencia fenoménica. Esto sugiere a su vez el principal problema que el argumento del dualismo de propiedades plantea al filósofo compatibilista: otorgar un mínimo grosor cognitivo a los conceptos fenoménicos que le permita dar cuenta del hiato explicativo pero que no le lleve a comprometerse con el fregeanismo mínimo.

Es de justicia reconocer que el propio Loar es consciente de este problema, pero lo que dice al respecto es a mi juicio demasiado poco:

Podemos suponer que el concepto corriente de dolor está relacionado con centros neurológicos del dolor; y que el concepto teórico de [oscilación córtico-talámica] está en cambio relacionado con centros verbales teóricos. Resulta poco sorprendente, en ese caso, que las relaciones neurológicas no permitan una conexión cognitiva directa” [Loar (2003), p. 117; traducción a mi cargo].

## VI. CONCLUSIONES

Tomar en consideración enunciados de reducción interteórica exige hacer algo que sólo está implícitamente supuesto en *Naming and Necessity*, a saber, indagar acerca de la semántica de los términos introducidos por teorías científicas. Pero cuando uno se pone a calibrar las opciones disponibles, descubre que la mayoría de ellas llevan a desactivar la versión Kripke-Chalmers del argumento dualista concebibilista. El dualista debe entonces encontrar justificación independiente para las pocas y algo rebuscadas opciones que no tienen esta consecuencia.

Pero del lado del fisicista compatibilista, que quiere por tanto respetar la tesis del hiato explicativo entre lo fenoménico y lo físico, también hay dificultades que superar. En la sección anterior he destacado dos especialmente importantes. En primer lugar, debe o bien desarrollar una epistemología modal racionalista que no se comprometa con la explicación kripkeana de las necesidades *a posteriori*; o bien defender una epistemología modal empirista.

En segundo lugar, debe también encontrar una teoría plausible de los conceptos fenoménicos que otorgue a éstos cierto grosor cognitivo capaz de explicar por qué existe supuestamente un hiato epistémico físico-fenoménico y que al mismo tiempo no sucumba al argumento del dualismo de propiedades\*.

*Departament de Filologia i Filosofia*  
*Facultat de Lletres*  
*Universitat de Girona*  
*Pl. Ferrater Mora 1, E-17071 Girona*  
*Email: david.pineda@udg.es*

#### NOTAS

\* La investigación en la que se enmarca este trabajo ha sido parcialmente financiada por el Ministerio de Educación y Ciencia del Gobierno de España mediante los proyectos de investigación HUM2007-61108 y HUM2005-04369. Desde aquí quisiera agradecer a dicha institución su apoyo y confianza.

#### REFERENCIAS BIBLIOGRÁFICAS.

- BLOCK, N. (2007), 'Max Black's Objection to Mind-Body Identity', en Alter y Walter (eds.), *Phenomenal Concepts and Phenomenal Knowledge*, Oxford University Press, pp. 249-306.
- BOYD, R. (1980), 'Materialism without Reductionism: What Physicalism Does Not Entail', en Block, N. (ed.): *Readings in the Philosophy of Psychology*, vol 1, Harvard University Press, pp. 67-106.
- CHALMERS, D. J. (2007), 'The Two-Dimensional Argument Against Materialism', de próxima aparición en Chalmers D. J., *The Character of Consciousness*, Oxford University Press. Versión electrónica en <http://consc.net/papers/2darguments.html>.
- FINE, K. (1994), 'Essence and Modality', *Philosophical Perspectives*, vol 8, pp. 1-16.
- JACKSON, F. (1982), 'Epiphenomenal Qualia', *Philosophical Quarterly*, vol. 32, pp. 127-36.
- KRIPKE, S. (1980), *Naming and Necessity*, Harvard University Press.
- LEVIN, J. (2002), 'Is Conceptual Analysis Needed for the Reduction of Qualitative States?', *Philosophy and Phenomenological Research*, vol. 64, pp. 571-91.
- LEVINE, J. (1983), 'Materialism and Qualia: the Explanatory Gap', *Pacific Philosophical Quarterly*, vol. 64, pp. 354-61.
- (1998), 'Conceivability and the Metaphysics of Mind', *Nous*, vol. 32, pp. 449-80.
- (2001), *Purple Haze. The Puzzle of Consciousness*, Oxford University Press.
- LOAR, B. (1997), 'Phenomenal States', en Block, Flanagan y Güzeldere (eds.), *The Nature of Consciousness*, Cambridge, Mss. MIT Press, pp. 597-616.
- (2003), 'Qualia, Properties, Modality', *Philosophical Issues*, vol 13, pp. 113-29.
- WHITE, S. (1983), 'Curse of the Qualia', *Synthese*, vol. 68, pp. 333-68.